Assessing Current & Future Infrastructure Hazards:
Forecasting Integrity using Machine Learning & Advanced Analytics

Lucy Romeo\textsuperscript{1,2}, Alec Dyer\textsuperscript{1,2}, Dakota Zaengle\textsuperscript{1,2}, Jake Nelson\textsuperscript{3}, Madison Wenzlick\textsuperscript{1,2}, Rodrigo Duran\textsuperscript{4}, Michael Sabbatino\textsuperscript{1,2}, Patrick Wingo\textsuperscript{1,2}, Aaron Barkhurst\textsuperscript{5}, Jennifer Bauer\textsuperscript{1}, and Kelly Rose\textsuperscript{1}

1. National Energy Technology Laboratory, Morgantown WV and Albany OR
2. Leidos Research Support Team, Morgantown WV and Albany OR
3. Oak Ridge Institute of Science and Technology, Austin TX
4. Theiss Research, San Diego CA
5. Mid-Atlantic Technology, Research & Innovation Center, Morgantown WV

Offshore Unconventional Resources Task 3, 2018-2021
Project Number DE FE-1022409
DOE-NETL's 2020 FE R&D Virtual Project Review Meeting – Oil & Natural Gas
October 2020
A few definitions to set the stage…

- **Artificial Intelligence**
  - “Programmed” intelligence

- **Machine Learning**
  - **Supervised ML**, the machine is trained, taught
  - **Unsupervised ML**, the machine learns on its own (Google’s cat video experiment)

- **Big Data**
  - Large volumes, variety, variability, velocity of data

- **Big Data Computing**
  - Computing engineering & systems to handle big data
What is the need?

- Rigs and platforms are designed for single use
- We are asking more from infrastructure
- Operations in offshore environments introduce hazards that can impact infrastructure integrity
- Need methods & models to assess existing infrastructure for future use (EOR, CS)

Project Objective

Execute intelligent analytics via an advanced analytical framework, to assess the current state of offshore infrastructure, evaluate infrastructure life, and identify technologies to reduce infrastructure hazards, costs, and extend infrastructure life.

Environmental conditions are getting more severe, need to prepare to prevent spills but also support response planning

Existing Platforms

as of February 2019

- Fixed
- Mobile
- Unknown
Approach & Results thus Far

• Build comprehensive dataset
• Perform data-driven analytics to evaluate infrastructure integrity
  1. Remaining lifespan
  2. Likelihood of future risk
• Apply data-driven advanced spatial, statistical, and Machine Learning (ML) models to quantify existing infrastructure integrity
• Release data and models through a smart, online platform hosted by Energy Data eXchange (EDX)

Values Delivered
• Identify & address critical hazards
• Reduce infrastructure hazards, costs
• Identify potential for extending infrastructure life for EOR

Meteorological & Oceanographic (MetOcean) impacts

Integrity risks

Offshore drilling regulator warns of bolt failures in Gulf of Mexico

Age-related infrastructure hazards

Data-Driven Approach & Driving Insights
Leveraging Big Data & Big Data Computing of the whole to inform the local

What we have:
- Structural information
- Incident records
- MetOcean data (Meteorological & Oceanographic data)

What we are incorporating:
- Geohazard data
- Historical incidents
Dataset Development

- Dataset includes:
  - Structural information
  - Structural- or weather-related incidents
  - Local environment history and anomalies

- Evaluated existing infrastructure integrity based on:
  - Location
  - Use
  - Design
  - Incident history

Initial analytics focusing on platforms:

**Sources:**
BOEMRE, BSEE, USCG, MMS

- **7,065** platform records
  - **1942-2020**
- **>4,000** incident records
  - **1956-2000, 2006-2018**

**Platform Type**
- Fixed
- Mobile Offshore Production Unit
- Unknown

**Fixed Platforms**
- Count: 1,736
  - Avg. Current Age (yrs): 35.6
- Count: 5,271
  - Avg. Age at Removal (yrs): 20.6
  - Std. Dev. Of Age at Removal (yrs): 13.4
  - Max. Age at Removal (yrs): 71.6
  - Min. Age at Removal (yrs): 0

**MOPU Platforms**
- Count: 50
  - Avg. Current Age (yrs): 14.5
- Count: 8
  - Avg. Age at Removal (yrs): 6.5
  - Std. Dev. Of Age at Removal (yrs): 2.5
  - Max. Age at Removal (yrs): 10.5
  - Min. Age at Removal (yrs): 2.6

---

MetOcean Data Extraction

>51,000 layers (>130 GB)

Sources include NOAA, US Navy, World Ocean Database, and external models

HYCOM

- Water velocity

WAVEWATCH III

- Wave height
- Wave direction
- Wave power
- Wind speed

IBTrACS Storms

- Storm characteristics

Proxies for corrosion

- Nitrate
- Dissolved oxygen
- Phosphate
- Temperature
- Silicate
- Salinity

Data example: Density of storm occurrences by wind speed

IBTrACS Storms

- NOAA WAVEWATCH III Hindcast and Reanalysis
- NOAA WAVEWATCH III Production Hindcast
- Expt. 50.1 Reanalysis
- Expt. 20.1
- Expt. 31.0
- Expt. 32.5

Platform Records

Exploratory Stats & Variable Analyses

- Key variable selection improved understanding of operational platform integrity
- Platforms in areas of more severe environmental conditions are removed on average 8 years sooner than platforms in less severe areas (Spatial autoregressive model)
- Applied AIC for model quality testing on > 40 variables

Strong relationship (0.95) between hurricanes and removal age (Pearson's Correlation)

Potential plateau in the number of times a platform can sustain extreme hurricane conditions (Ordinary Least Squares regression)
Machine Learning & Advanced Algorithms

Applied multiple methods with comprehensive dataset to model existing platform integrity:

• Predicting Lifespan
  1. Gradient Boosted Classifier
  2. Artificial Neural Network (NN) Classifier
  3. Geographically Weighted Regression

• Assessing Risk Likelihood
  1. Gradient Boosted Regression
  2. NN Regression

Analyses through ML & advanced statistical models

With new information, model predictions improve

Benefit of Multiple Models

1. Evaluated available data
2. Identified key parameters
3. Assessed multiple approaches
4. Compare results for internal model validation

**Gradient Boosting**
- Stagewise additive ML model that optimizes model loss by adding weak learners (i.e. decision trees)

  **Strengths**
  - Handles outliers and able to perform when data is missing

  **Limitations**
  - Difficult to scale model
  - Requires careful tuning of parameters

**NN**
- Deep learning ML model where data are passed through multiple layers that learn weights and biases, which are combined and transformed to learn the phenomena being modeled.

  **Strengths**
  - Handles complex and non-linear relationships

  **Limitations**
  - Easy to overfit
  - Difficult to interpret
  - Many parameters to tune

**Geographically Weighted Regression**
- Spatial model that fits regression equations to each feature (platform), based on dependent and explanatory variables of other local features

  **Strengths**
  - Captures local variation in spatial data
  - Can be used for prediction

  **Limitations**
  - Computational overhead increases with data volume
  - Local collinearity must be considered
  - Assumes linear relationships

---

Predicting remaining lifespan

**Gradient Boosted Classification**
85-89% accuracy

**Artificial Neural Network Classification**
81-86% accuracy

Accuracy = \( \frac{\text{Correct Predictions}}{\text{Total Predictions}} \)

Running *multiple models* allows us to better understand and *internally validate results*.

Accuracy will *increase* by giving the model *more accurate information* to learn from.
Geographically Weighted Regression
Predicting current age based on health & environment of existing infrastructure

Top Parameters
Mean surface salinity
Distance to shore
Category 4 hurricane days
Logged max. wave power
Max. wave period

0.85 correlation between known and predicted age

Explains 51 to 97% of the variance in the data

Results from model conclude age of removal has spatial nonstationarity

Gradient Boosted Regression

- **85-99% accuracy**, based on cross-validation method with 5 folds to evaluate average model performance

NN Regression

- Negative accuracy – Anticipating better fitting model with additional data and hyper tuning

**Currently resurrecting and incorporating historical incidents (1956-2000)**

**Expecting models to evolve with new data**

**Top Parameters**

- Number of structural incidents
- Number of incidents
- Number of incidents during production operations
- Number of electrical incidents
- Category 5 storm (number of days)

- Maximum wind velocity
- Wave height (25th percentile)
- Maximum reported wind gust
- Wind magnitude (50th percentile)
- Number of incidents during motor vessel operations

- Injection code (injecting gas or water)
- Commingling production
- Allocation Meter flag (platform has allocation meter)
- Production Equipment Flag (installed production equipment)
- Authority status

**Training R²**

- ≤0.44
- ≤1.0

**Testing R²**

Increasing Access to Data & Models
Through NETL’s Energy Data eXchange (EDX)

- Integrate ML, big data, and analytical outputs with online platform
- Leverage award-winning Offshore Risk Modeling Suite
- Transform models into online tools to perform real-time analyses to better understand offshore infrastructure integrity
- Release tools and platform through EDX

Initial release via EDX for select members of DOE/NETL community at end of EY20

https://edx.netl.doe.gov/offshore
Intelligent Offshore Infrastructure Integrity Analyses
Timeline (present to March 2021)

**Next steps**
- Complete filling data gaps
- Refine & validate models
- Publish technical report on data and methods (Nelson et al., *in progress*)

**Upcoming Milestones**

<table>
<thead>
<tr>
<th>Date</th>
<th>Description</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>12/20</td>
<td>Draft promotional items for upcoming release of online advanced infrastructure analytical platform, Version 2.</td>
<td>On track</td>
</tr>
<tr>
<td>02/21</td>
<td>Submit article on intelligent analytics for infrastructure hazards and development.</td>
<td>On track</td>
</tr>
</tbody>
</table>

**Upcoming Deliverables**

<table>
<thead>
<tr>
<th>Date</th>
<th>Description</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>12/20</td>
<td>Submit technical report detailing development of intelligent analytics.</td>
<td>On track</td>
</tr>
<tr>
<td>03/21</td>
<td>Advanced release of online interactive analytical platform via EDX.</td>
<td>On track</td>
</tr>
</tbody>
</table>

https://edx.netl.doe.gov/offshore/portfolio_items/assessing_current_and_future_infrastructure_hazards/
Upcoming Publications


Upcoming Presentations


Past Presentations


Key Takeaways

• Built comprehensive infrastructure dataset
• Developed novel ML analytical models to predict existing lifespan and risk
• Continuing to add data to make models smarter
• Pubs are in prep
• Integration of data, models, and tools on virtual platform

Values Delivered

Understand existing infrastructure integrity
Identify potential for extending infrastructure life for energy security
Minimize cost, maximize safety

Lucy.Romeo@netl.doe.gov
Kelly.Rose@netl.doe.gov

Disclaimer & Acknowledgement

• **Disclaimer:** This presentation was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference therein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed therein do not necessarily state or reflect those of the United States Government or any agency thereof.

• **Acknowledgement:** Parts of this technical effort were performed in support of the National Energy Technology Laboratory’s ongoing research under the Offshore Unconventional Resources – DE FE-1022409 by NETL’s Research and Innovation Center, including work performed by Leidos Research Support Team staff under the RSS contract 89243318CFE000003.